Understanding the heterogeneity of high-grade CIN lesions

Bierkens, M.

2013

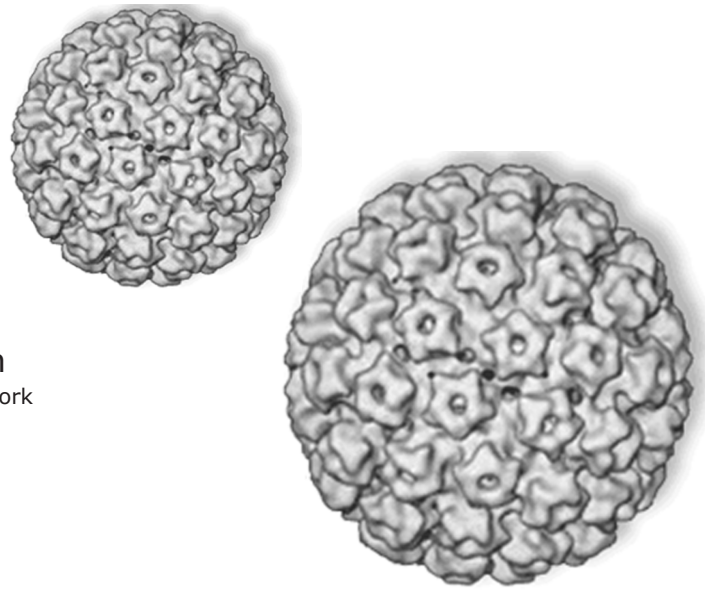
document version
Publisher's PDF, also known as Version of record

Chapter 4

Focal aberrations indicate EYA2 and hsa-miR-375 as oncogene and tumour suppressor in cervical carcinogenesis

Genes, Chromosomes Cancer 2013 Jan;52(1):56-68

Mariska Bierkens*
Oscar Krijgsman*
Saskia M. Wilting
Leontien Bosch
Annelieke Jaspers
Gerrit A. Meijer
Chris J.L.M. Meijer
Peter J.F. Snijders
Bauke Ylstra
Renske D.M. Steenbergen
*authors contributed equally to this work



Abstract

Cervical cancer results from persistent infection with high-risk human papillomavirus (hrHPV). Common genetic aberrations in cervical (pre)cancers encompass large genomic regions with numerous genes, hampering identification of driver genes. This study aimed to identify genes functionally involved in HPV-mediated transformation by analysis of focal aberrations (<3 Mb) in high-grade cervical intraepithelial neoplasia (hgCIN).

Focal chromosomal aberrations were determined in high-resolution arrayCGH data of 60 hgCIN. Genes located within focal aberrations were validated using two external gene expression datasets or qRT-PCR. Functional roles of candidate genes EYA2 (20q13) and hsa-miR-375 (2q35) were studied by siRNA-mediated knockdown and overexpression, respectively, in hrHPV-containing cell lines.

We identified 74 focal aberrations encoding 305 genes. Concurrent altered expression in hgCIN and/or cervical carcinomas compared with normal cervical samples was shown for ATP13A3, HES1, OPA1, HRASLS, EYA2, ZMYND8, APOBEC2 and NCR2. Gene silencing of EYA2 significantly reduced viability, migratory capacity and anchorage-independent growth of HPV16-transformed keratinocytes. For hsa-miR-375, a direct correlation between a (focal) loss and significantly reduced expression was found. Downregulation of hsa-miR-375 expression was confirmed in an independent series of cervical tissues. Ectopic expression of hsa-miR-375 in two cervical carcinoma cell lines reduced cellular viability.

Our data provide a proof of concept that chromosomal aberrations are actively contributing to HPV-induced carcinogenesis and identify EYA2 and hsa-miR-375 as oncogene and tumour suppressor gene, respectively.

Introduction

Persistent infection with high-risk human papillomavirus (hrHPV) has been causally related to cervical cancer development.[1] Squamous cell carcinomas (SCC) are preceded by precursor lesions, called cervical intraepithelial neoplasia (CIN).[2] Based on histology CIN can be classified as low-grade (CIN1) or high-grade (CIN2/3). Low-grade CIN are mainly associated with productive viral infections.[3] High-grade CIN (hgCIN) usually harbour transforming HPV infections and are characterised by deregulated expression of viral oncogenes E6 and E7 in proliferating cells and diffuse CDKN2A overexpression.[3,4] Overexpression of the viral oncogenes leads to accumulation of (epi)genetic changes in the host cell genome and may drive progression to cervical cancer.[1] Array comparative genomic hybridisation (arrayCGH) on hgCIN has revealed common chromosomal aberrations, e.g. gain of 3q.[5-10] Though these aberrations are in part lesion, tumour, or HPV type related[11-13], it has not been unequivocally demonstrated if they actively contribute to HPV-induced carcinogenesis.

Each tumour genome harbours a mixture of genetic aberrations with genes directly responsible for its development (drivers) and random events (passengers). Distinguishing driver from passenger aberrations is crucial for understanding cervical cancer development and will aid in the identification of functional biomarkers or even molecular targets for therapeutic intervention. Identification of driver genes has been a major challenge in early (array)CGH studies due to the large size of detected aberrations and the numerous genes located therein[11], making functional validation of potential candidates virtually impossible. Additionally, cancer genomes generally display a high degree of chromosomal instability, including many passenger aberrations.[14] The use of premalignant lesions aids in the identification of driver genes due to a lower degree of genetic chaos.

Increased resolution of arrayCGH platforms and their applicability to formalin-fixed paraffin-embedded (FFPE) tissues have yielded new insights into small chromosomal aberrations of only several Mb previously undetected or erroneously classified as outliers.[15] Small somatic aberrations, so-called focal aberrations (commonly regarded as <3 Mb) harbour only a small number of genes and thereby aid the identification of drivers in a cancer genome. Recent studies showed focal aberrations to be enriched for cancer genes and genes frequently mutated in cancer.[16-18] Analysis of focal aberrations and genes therein by combined bio-informatical analysis and experimental biology has led to the discovery of novel tumour suppressor genes and oncogenes in various non-virally induced epithelial tumours such as lung, colon and breast cancer, including CDH20, PRKAA2, GPR124, ARFRP1 and hsa-miR-101.[16-19] The role and order of chromosomal aberrations in the genetic history of these epithelial tumours is unclear. This is different for cervical cancer where persistent infection with hrHPV may provoke tumour development by inducing genetic instability, chronologically succeeded by selected (epi)genetic changes providing a growth advantage, that yet become evident in hgCIN. In this study we aimed to identify potential driver genes of HPV-induced

cervical carcinogenesis by analysis of focal aberrations in 60 hgCIN.

Materials and Methods

Tissue samples and cell lines

Previously, 60 HPV-positive, formalin-fixed paraffin-embedded (FFPE) hgCIN (of which 56 CIN3 and 4 CIN2) were analysed by high-resolution comparative genomic hybridisation microarray (arrayCGH).[10,13] HPV types of CIN3 were: HPV16 n=25, HPV18 n=4, HPV31 n=14, HPV33 n=2, HPV45 n=3, HPV51 n=2, HPV52 n=3, HPV58 n=2 and HPV69 n=1; of CIN2: HPV16 n=1, HPV31 n=2 and HPV45 n=1. Samples were compared with pooled DNA from FFPE cervical biopsies of 5 women diagnosed with leiyoma, but with no cervical disease. These biopsies were negative for the presence of either low- or high-risk HPV.

All lesions analysed showed strong diffuse staining for CDKN2A, indicative of a transforming HPV-infection[4] and were obtained from the population-based screening study Amsterdam (POBASCAM)[20]. An additional set of HPV-positive frozen biopsies of 6 normal cervical squamous epithelial specimens, 13 CDKN2A-positive hgCIN and 9 SCC was used to validate altered expression of hsa-miR-375. These specimens were collected during routine clinical practice at the Department of Gynaecology and Obstetrics and stored at the Department of Pathology of the VU University Medical Center (Amsterdam, The Netherlands).

The HPV16-immortalised keratinocyte cell line FK16A was established and cultured as described previously.[21] Human cervical carcinoma cell lines SiHa and CaSki were obtained from the American Type Culture Collection (Manassas, VA) and cultured as described previously.[22] All cell lines were at around the time of functional experiments tested for the presence of the correct HPV type, the presence of known copy number and DNA methylation alterations for authentication of the cell lines.

ArrayCGH procedures and data analysis of potential focal aberrations

Microdissection and DNA extraction of the hgCIN was performed as described previously.[10,13] Whole genome amplification using the Bioscore kit (Enzo Bioscore™ Screening and Amplification, Enzo Life Sciences, Farmingdale, NY) was performed as described by Buffart *et al*.[23] Labeling of DNA with Cy3 or Cy5 was performed with the Enzo Genomic DNA labeling kit according to the manufacturer's instructions (Enzo Life Sciences) using 500 ng of amplified DNA. Hybridisations were performed on slides containing two 105K arrays. Each array contained 99,000 synthetic 60-mer oligonucleotides, with an overall median probe spacing of 21.7 kb (Agilent Technologies, Palo Alto, CA). Across array CGH (aaCGH) as well as across slide comparison was used, in which a sample was either hybridised to a reference, or to another sample as described by Buffart *et al*.[24] Dye channels were digitally swapped in order to compare each sample with a reference. Data extraction and calling of gains and losses was performed as described previously.[10] Array data are

available from the Gene Expression Omnibus (GEO, http: //www.ncbi.nlm.nih.gov/projects/geo/) through series accession number GSE34575.

Downstream analysis for focal aberrations was performed and plots were made using the statistical programming language R version 2.11.1. Log_2-ratios of signal intensities between hgCIN and reference for every probe were median-normalised and post-segmentation mode normalisation was performed.[25] To discriminate between somatic focal aberrations and germ-line copy number variants (CNVs) all aberrations smaller than 3 Mb were identified using the processed copy number data and compared with a database of known CNVs in the healthy population using the database of genomic variants (DGV).

Only gains and losses detected at least twice (recurrent) in the dataset were considered. These recurrent focal aberrations were further filtered for aberrations that were only affected in a single direction (gain or loss). For each recurrent focal aberration the smallest genomic overlap of the focal aberrations and the frequency in the dataset was determined (the high frequency region, HFR). Genes and miRNAs were retrieved using biomaRt (R/Bioconductor) and Ensembl (hg18/NCBI 36, ensemble 54). Potential driver genes were assumed to be located completely within the HFR in case of both gains and losses and were selected based on the total frequency of occurrence in the data series.

Expression analysis by quantitative reverse-transcription PCR (qRT-PCR)

Expression of EYA2 was measured on the ABI7500 Fast Real-Time PCR System (Applied Biosystems) using primers and probe described in Farabaugh *et al*.[26] The house keeping gene snRNP U1A was included for normalisation purposes, as described by Henken *et al.*[27] Expression values were normalised to the reference using the comparative C_t method ($2\text{-}\Delta C_t$)[28]. qRT-PCRs were performed using Universal PCR master mix (Applied Biosystems).

Expression of hsa-miR-375 was measured using TaqMan microRNA assays following the manufacturer's instructions (Assay ID: 000564; Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands) on the ABI7500 Fast Real-Time PCR System. The small nucleolar RNA transcript RNU43 (Assay ID 001095; Applied Biosystems) was included as internal reference for hsa-miR-375 expression. Expression values were normalised to the reference using the comparative C_t method ($2\text{-}\Delta C_t$).[28]

Validation of candidate driver genes in independent expression array studies

Two independent single-channel, expression microarray datasets (Affymetrix Human Genome U133A Array, HG-U133A) were used to validate the candidate genes in the most frequently occurring focal gains and losses. Normalised expression data were downloaded from the GEO database: GSE7803 and GSE9750.[29,30] From GSE7803, 38 samples were used, including 10 normal cervical samples, 7 hgCIN and 21 SCC. From GSE9750, 57 samples were used, including 24 normal cervical samples and 33 SCC. Probes were annotated using Ensembl (ensembl.org) based on hg18/NCBI36 (ensembl 54).

siRNA and cDNA transfection

Transfections were performed using Dharmafect2 (Dharmacon, Lafayette, LA) and Lipofectamine 2000 (Invitrogen) for siRNA and cDNA, respectively, according to the manufacturer's recommendations.

FK16A cells, passage 255, were used for experiments using siRNA targeting either EYA2 'ON-TARGETplus SMARTpool EYA2' (Cat#L-017233-00-0005, Dharmacon, Lafayette, USA) or non-targeting siRNA pool 'ON-TARGETplus Non-targeting Pool' (Cat#D-001810-10-05, Dharmacon) according to the manufacturer's instructions. As positive control for transfection efficiency, PLK1 specific siRNAs 'ON-TARGETplus SMARTpool PLK1' (Cat#L-003290-00-0005) was used.

SiHa and CaSki cells were transfected with either hsa-miR-375 construct pcDNA4/H1/miR-375 or empty vector pcDNA4/H1 (a generous gift from dr. R.J. Perera[31]).

After transfection, silencing (EYA2) and ectopic overexpression (hsa-miR-375), respectively, were checked by qRT-PCR as described above.

Cell viability analysis

Cell viability was measured using a colorimetric (MTT-tetrazolium) assay (ICN Biomedicals Inc, Cleveland, OH) according to the manufacturer's instructions. Cells were transfected with the respective siRNA and cDNA constructs and equal cell numbers for all conditions were seeded in triplicate in 96-well plates. MTT conversion was measured at days 0 and 2. Cellular viability was calculated for all tested conditions by subtracting the measurement of day 0 from day 2. All experiments were performed in duplicate, starting from independent transfections.

Migration and anchorage-independent growth analysis

Migration and anchorage-independent growth was determined as described previously.[27] In brief, for migration analysis, siRNA transfected cells were plated at high confluency and uniformly scratched to create a cell-free gap. Plates were photographed again after 30h.

To examine anchorage-independent growth, 5000 siRNA transfected cells were plated in semi-solid agarose. Colonies were photographed and counted after 3 weeks of incubation.

Statistical analysis

To test whether focal aberrations were significantly enriched for cancer related genes as published in the cancer census list[32], enrichment analysis was performed. Ten thousand sets of random aberrations were generated of the same size as the focal aberrations identified in our study and genes therein were matched to genes of the cancer census list. Overlap of genes for each random set with genes of the cancer census list was calculated and expressed as a p-value.[18]

Gene expression between histological groups was assessed using the Mann-Whitney U test for both qRT-PCR and microarray data. Association between the number of gains and losses, exclusively gained or lost, and CNVs or somatic focal

aberrations was determined using a X^2-test. P-values smaller than 0.05 were considered statistically significant.

Results

Focal aberrations in hgCIN

A total of 235 recurrent focal aberrations (<3 Mb) were detected in 60 hgCIN. The frequency with which these chromosomal gains and losses were detected is shown in Fig. 1. Overlap with known germ-line copy number variants (CNVs) in the healthy population marked 138 of these aberrations as CNVs.[33] Differences were observed between gains and losses overlapping known CNVs and remaining potential somatic aberrations. For example, 84% of CNVs versus 97% of potential somatic aberrations contained a gene. There was a significant difference in the number of aberrations exclusively affected as either a gain or a loss between CNVs and somatic aberrations, i.e. 42% and 76%, respectively (X^2-test, $p<0.01$). Therefore, all focal aberrations showing both gains and losses in the dataset were considered potential CNVs and excluded from further analysis. This reduced the list of somatic focal aberrations to 74 genomic regions, indicated in black in Fig. 1 and listed in Supplementary Table I. The most frequent aberrations and genes therein are summarised in Table 1. For each recurrent focal aberration the high frequency region (HFR) was identified. A total of 305 genes were located in these HFRs (range: 0-28 genes per focal aberration). Focal aberrations were significantly enriched for known cancer related genes (enrichment analysis: $p=0.02$) based on the Cancer Gene Census database (http: //www.sanger.ac.uk/genetics/CGP/Census), including ABL1, ARHGEF12, CDX2, ELN, EML4, FEV, FLT3, FNBP1, HOXD11 and JAK2 (Supplementary Table I).

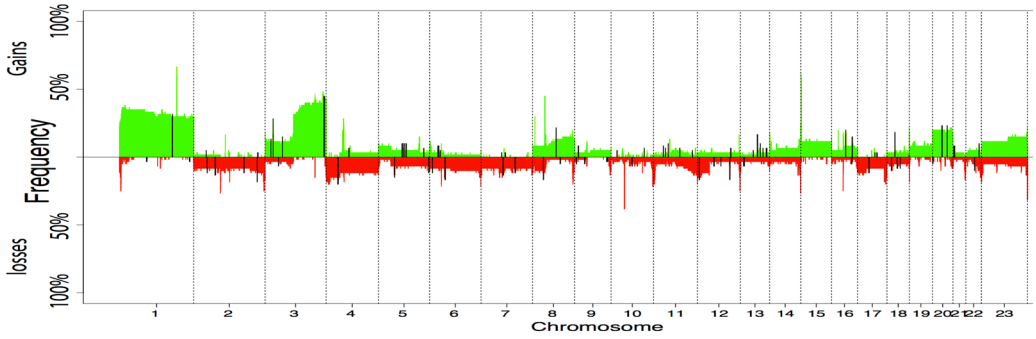


Figure 1: Frequency plot of all copy number aberrations and focal aberrations. Percentages of gains (positive Y-axis) and losses (negative Y-axis) for each oligonucleotide are shown for the 60 hgCIN. Focal aberrations are indicated in black.

Table 1: Summary of the most frequent focal aberrations (exclusively affected as a gain or a loss) detected and associated full-length genes therein.

					Focal[a]		Total[b]		
	Cytoband	Start	End	Mb	Gain	Loss	Gain	Loss	Genes within HFR[c]
Gain									
	20q13.12	44945757	45671015	0.73	4	0	14	0	ZMYND8; EYA2
	3q29	194023770	195677189	1.65	2	0	27	0	ATP13A3; GP5; LRRC15; CPN2; HES1; OPA1; ATP13A4; ATP13A5; HRASLS
	1q25.2	177742624	178392811	0.65	2	0	17	0	TOR1AIP1; CEP350; TDRD5; FAM163A; TOR1AIP2; NPHS2
	20q11.22	31552311	31700561	0.15	2	0	14	0	CBFA2T2
Loss									
	2q35	219554248	219604836	0.05	0	5	0	10	CRYBA2; FEV; hsa-miR-375
	6p21.1	40624689	41748197	1.12	0	5	0	10	MDFI; FOXP4; NCR2; TREM1; TREML4; TREML3; TREML2P; TREML2; TREM2; TREML1; NFYA; C6orf130; APOBEC2; BZRPL1; UNC5CL
	4p13	41832776	42663376	0.83	0	3	0	12	ATP8A1; SHISA3
	7q11.23	73105632	73429720	0.32	0	3	0	8	RFC2; LAT2; EIF4H; LIMK1; hsa-miR-590
	2q31.1	176673584	176745707	0.07	0	3	0	7	hsa-miR-10B; HOXD3; HOXD4; HOXD8; HOXD9; HOXD10; HOXD11; HOXD12
	7q31.1	108206652	109055550	0.85	0	2	0	7	C7orf66
	17q11.2-q12	28368756	29488485	1.12	0	2	0	7	ACCN1
	10q22.1	72671625	73417685	0.75	0	2	0	4	PSAP; C10orf54; C10orf105; CDH23; SLC29A3

[a]The number of samples found to have a focal aberrations.

[b]The total number of samples in which this genomic location was altered.

[c]All genes that are located full-gene within the high frequency region (HFR) of the focal aberration.

Expression of protein-coding genes within focal aberrations

Expression of genes located within the 2 most frequent focal gains (20q13 and 3q29) and losses (2q35 and 6p21.1) (Table 1) was examined in 2 independent cervical cancer mRNA expression microarray datasets (GSE7803 & GSE9750).[29,30] The mRNA expression of genes within the 4 focal aberrations, showing altered expression concurrent with the chromosomal aberration in either one or both datasets, was plotted per histological grade (Fig. 2).

Focal gain at 20q13 was detected 4 times and the genomic region was gained in 10 other samples as part of a larger aberration. It harboured 2 genes: ZMYND8 and EYA2. In the first dataset, both genes showed increased expression in hgCIN and SCC versus normal. Increased expression was only significant for ZMYND8 in SCC versus normal in the second dataset ($p=0.03$). Increase of EYA2 expression nearly reached significance for hgCIN versus normal in the first dataset ($p=0.06$) and for SCC versus normal in the second dataset ($p=0.06$). Focal gain at 3q29 was detected 2 times and the genomic region was gained in 25 other samples as part of a larger aberration. Expression data in the external datasets was available for

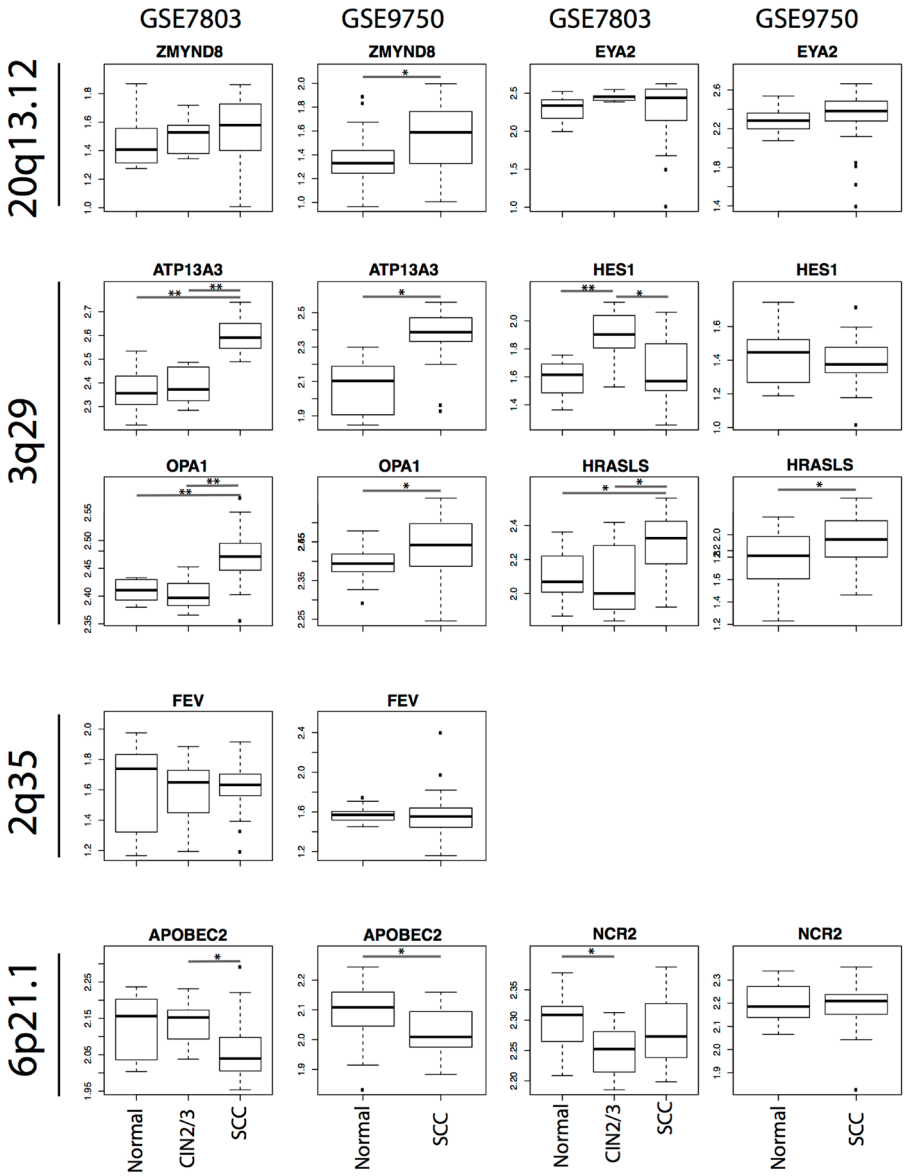


Figure 2. Expression of candidate driver genes. Boxplots showing mRNA expression of genes located within the 2 most common focal gains: 20q13 (ZMYND8 and EYA2) and 3q29 (ATP13A3, HES1, OPA1 and HRASLS) and the 2 most common focal losses: 2q35 (FEV) and 6p21 (APOBEC2 and NRC2) as obtained from 2 external datasets: left GSE7803; right GSE9750. For dataset GSE7803 expression data were available for normal cervix (n=10), hgCIN (CIN2/3, n=7) and squamous cell carcinoma (SCC, n=21). For dataset GSE9750 expression data were available for normal cervix (n=24) and SCC (n=33). *p<0.05, **p<0.01.

ATP13A3, GP5, LRRC15, CPN2, HES1, OPA1 and HRASLS, of which ATP13A3, HES1, OPA1 and HRASLS showed altered expression in SCC and/or hgCIN consistent with the copy number data. Expression of ATP13A3 was significantly higher in SCC versus normal in both datasets (both $p<0.001$) and in SCC versus hgCIN ($p<0.001$). Expression of HES1 was significantly higher in hgCIN versus normal ($p=0.003$), but there was no differential expression between SCC and normal in either dataset. OPA1 expression was significantly increased in SCC versus normal in both datasets ($p<0.001$ and $p=0.02$, respectively), yet expression in hgCIN and normal was similar. HRASLS was significantly higher expressed in SCC versus normal ($p=0.02$ and $p=0.04$) as well as SCC versus hgCIN ($p=0.04$). For the other genes at 3q29, ATP13A4 and ATP13A5, no expression data were available.

Focal losses at both 2q35 and 6p21.1 were detected 5 times and the genomic region was lost in 5 other samples as part of a larger aberration. Expression of FEV (2q35) was slightly decreased in hgCIN and SCC compared with normal. No (reliable) data were available for the genes, CRYBA2 and hsa-miR-375. Expression of APOBEC2 (6p21.1) was slightly decreased in hgCIN versus normal and was (significantly) lower in SCC compared with normal in both datasets ($p=0.07$ and $p=0.02$, respectively). Expression of NCR2 (6p21.1) was significantly reduced in hgCIN versus normal ($p=0.04$), but not in SCC versus normal in both datasets. The remaining genes located at 6p21.1 either showed no (concurrent) change in expression (TREM1, TREML2, NFYA and BZRPL1) or were absent in the datasets (MDFI, FOXP4, TREML4, TREML3, TREML2P, TREM2, TREML1, C6orf130 and UNC5CL).

Increased EYA2 expression associated with focal gain at 20q13 is functionally relevant

To investigate the potentially functional relevance of 20q13 focal gain in cervical carcinogenesis, EYA2 expression was silenced in late passage (p255) HPV16-immortalised keratinocytes (FK16A) using siRNA. These cells are characterised by both copy number gain of 20q[34] and increased EYA2 expression compared with primary foreskin keratinocytes (Supplementary Fig. I). Silencing of EYA2 expression was confirmed by qRT-PCR (Fig. 3A). Cell viability was significantly reduced in cells transfected with EYA2 siRNA compared with non-targeting siRNA transfected cells (Fig. 3B). Cells transfected with EYA2 siRNA showed reduced migratory capacity and significantly reduced anchorage independent growth compared with non-targeting siRNA transfected cells (Fig. 3C-E).

Reduced expression of hsa-miR-375 associated with focal loss at 2q35 is functionally relevant

The most frequent focal loss at 2q35 was detected in 5 samples and as a larger aberration in 5 extra samples (Table 1). The HFR contained CRYBA2, FEV and hsa-miR-375 (Fig. 4). Expression of FEV was not significantly decreased in the external datasets (Fig. 2). Though no (reliable) expression data were available for CRYBA2 or hsa-miR-375, we further investigated hsa-miR-375 since the stability of miRNAs allowed us to determine hsa-miR-375 expression in a subset of the FFPE hgCIN

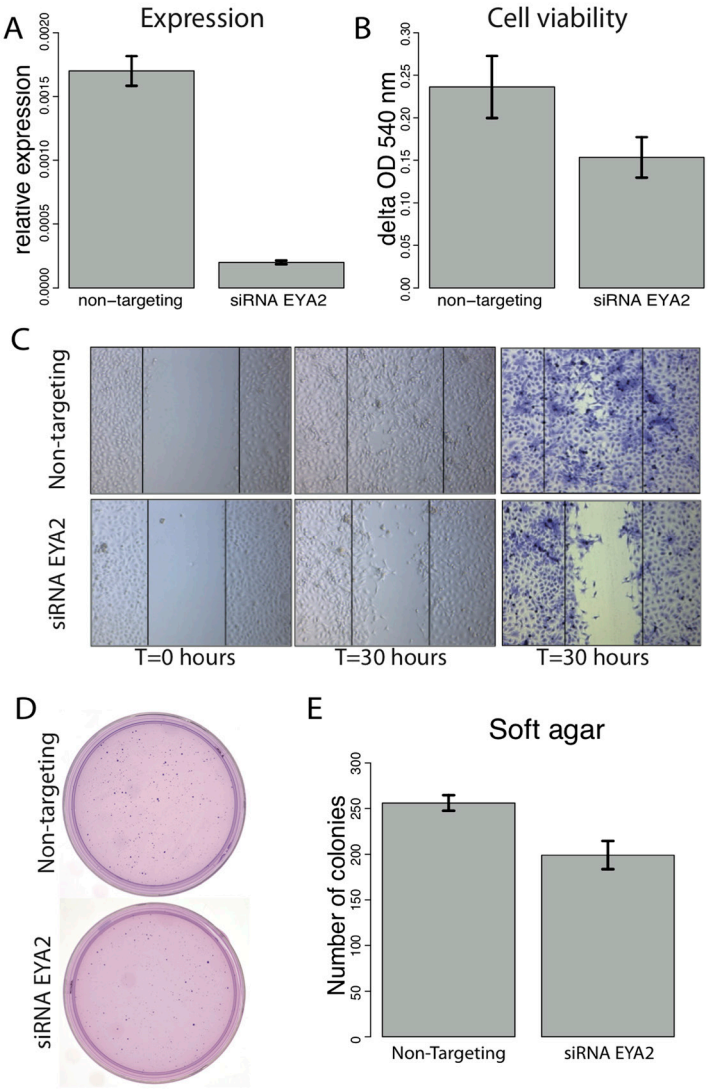


Figure 3: The effect of siRNA mediated silencing of EYA2 in FK16A cells. (A) Relative EYA2 mRNA expression levels in FK16A cells transfected with non-targeting siRNAs (left) and with a pool of EYA2-specific siRNAs (right), showing significantly reduced EYA2 mRNA expression in EYA2 siRNA transfectants. (B) Silencing of EYA2 significantly decreased cell viability. (C) Results of the scratch assay indicate decreased migratory capacity in cells transfected with EYA2 siRNAs. Photographs in left panel are taken immediately after scratching (t=0 hours). Middle and right panel shows same cells after 30h of incubation, the right panel being stained with crystal violet. (D) Results of anchorage-independent growth assay indicate decreased colony formation ability of cells transfected with EYA2 siRNAs. (E) Quantification of anchorage-independent growth assay.

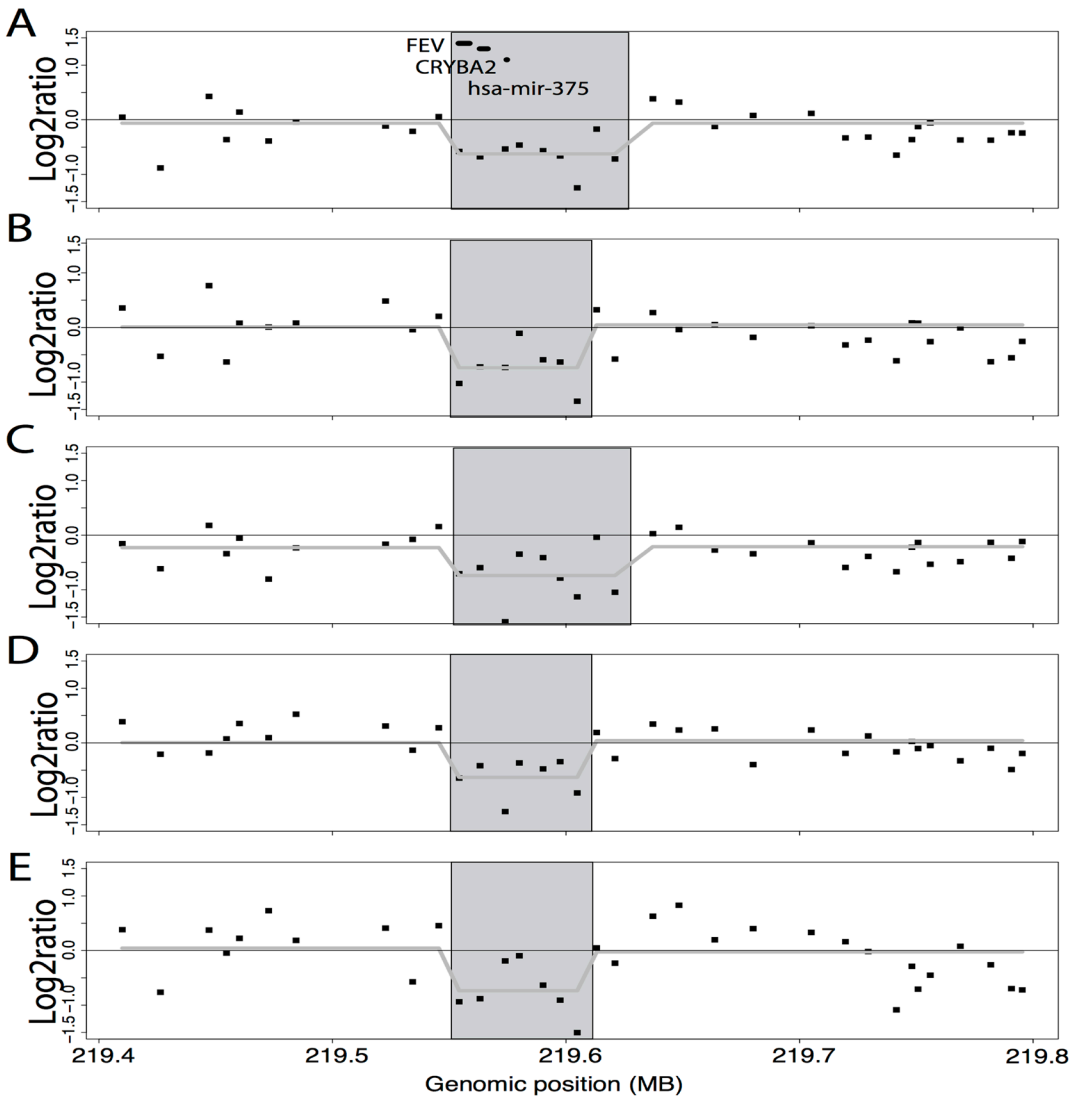


Figure 4: Focal loss at 2q35. Schematic representation of the focal deletion at chromosome 2q35 as detected in 5 of 60 hgCIN (A-E). X-axis represent the $\log_2$-ratio of the probes, Y-axis the genomic location. The grey line indicates the segmented values as obtained using circular binary segmentation (CBS). Bars at the top represent the exact location of the genes. Deleted genomic regions are indicated with grey fields. All samples, except C, show a focal deletion in an otherwise unaffected region; (C) shows a focal deletion within a larger loss at 2q35.

used in the arrayCGH experiments. Hsa-miR-375 expression could be measured in 4 out of 5 samples with focal loss, 5 samples with a larger loss and in 28 samples with no aberration at that locus. Hsa-miR-375 expression was significantly lower in hgCIN with (focal) loss compared with lesions with no aberration (p=0.04, Mann-Whitney U test; Fig. 5A). Reduced hsa-miR-375 expression was verified in an independent series of samples consisting of frozen specimens of 6 normal cervices, 13 hgCIN and 9 SCC (Fig. 5B). Expression of hsa-miR-375 decreased in a step-wise pattern from normal cervical epithelium to hgCIN to SCC (SCC compared with normal: p=0.003). In order to investigate the effect of hsa-miR-375 on cell viability, SiHa and CaSki cells were transiently transfected with either an empty vector or a hsa-miR-375 expression vector. Ectopic expression of hsa-miR-375 was

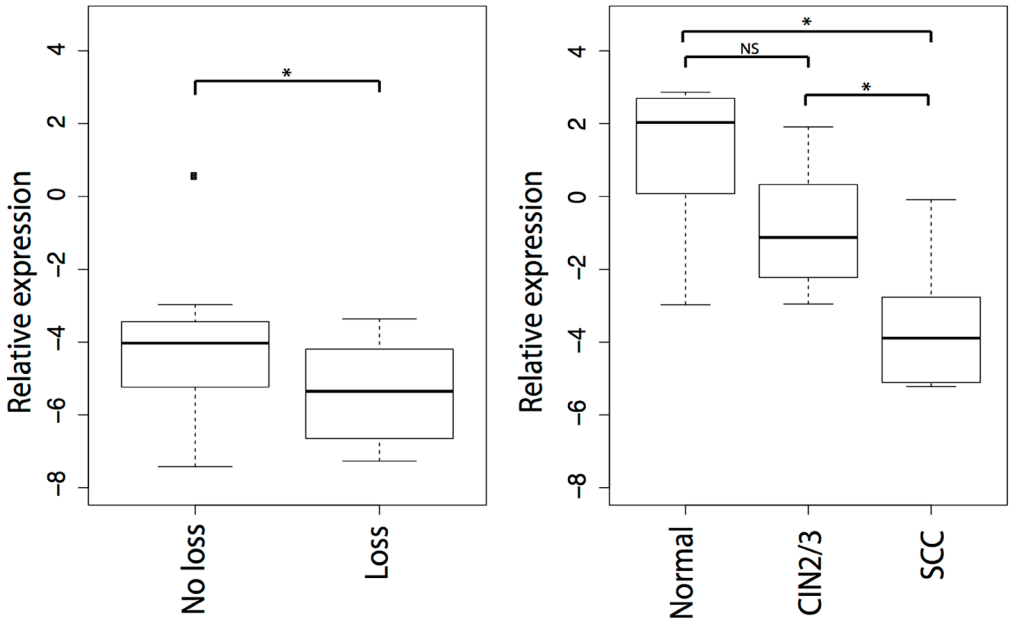


Figure 5: Hsa-miR-375 expression in cervical tissue samples. (A) HgCIN with a (focal) loss of 2q35 show significantly decreased hsa-miR-375 expression compared with lesions with a normal copy number. (B) Relative hsa-miR-375 expression in an independent series of cervical specimens (6 normal cervices, 13 hgCIN and 9 SCC). Hsa-miR-375 expression is significantly decreased in SCC compared with normal cervical samples as well as hgCIN. *p<0.05, NS is not significant.

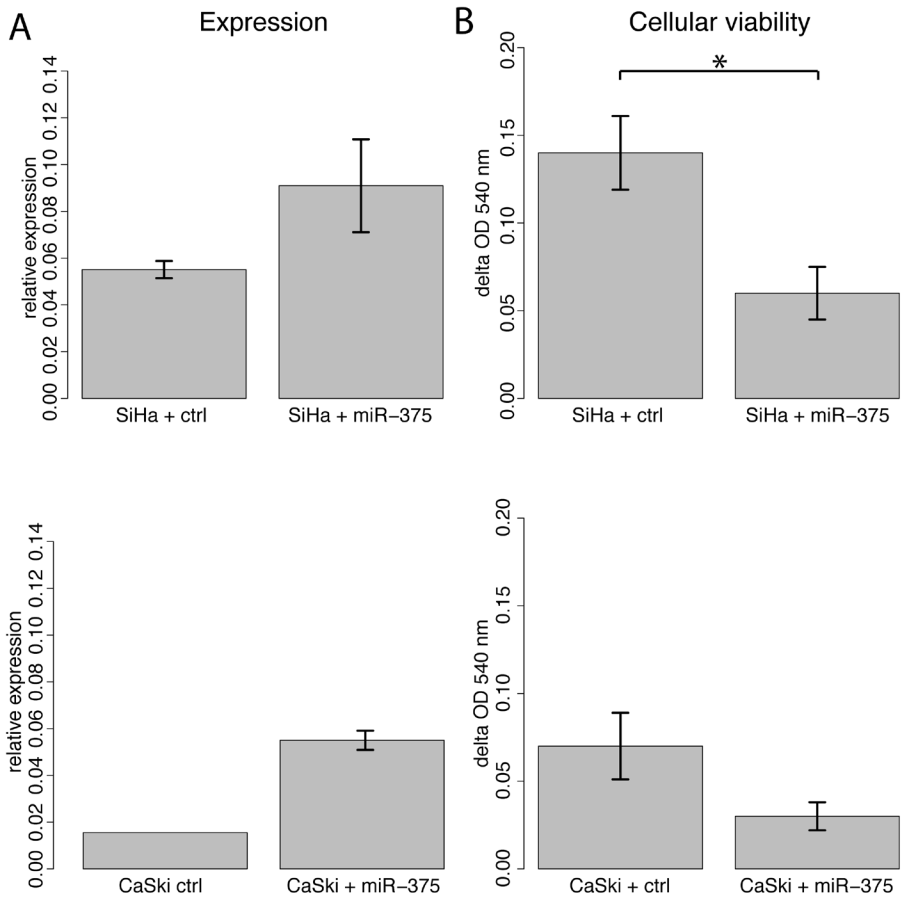


Figure 6: Ectopic expression of hsa-miR-375 in SiHa and CaSki cells. (A) SiHa and CaSki cells transiently transfected with hsa-miR-375 show significantly increased hsa-miR-375 expression compared to cells transfected with empty vector. (B) Ectopic of hsa-miR-375 resulted in decreased cellular viability. *$p<0.05$, NS is not significant.

confirmed by qRT-PCR (Fig. 6A). Cell viability was significantly reduced in SiHa cells transfected with hsa-miR-375 compared to the empty vector control (t-test, $p<0.05$; Fig. 6B). Cell viability was also reduced in CaSki cells upon ectopic hsa-miR-375 expression, yet this did not reach significance (Fig. 6B).

Discussion

In early stages of cervical cancer development we identified 74 recurrent focal aberrations encoding 305 genes using high-resolution arrayCGH data of 60 hgCIN. These focal aberrations were significantly enriched for cancer census genes. Using 2 independent mRNA expression datasets a number of genes located within the focal aberrations were identified as candidate oncogenes and tumour suppressors, including ZMYND8 and EYA2 (20q13). Additionally, hsa-miR-375 (2q35) was acknowledged as a candidate tumour suppressor gene by expression analysis of 2 series of cervical specimens. Functional validation studies on EYA2 and hsa-miR-375 provided proof of concept that chromosomal aberrations are actively contributing to HPV-induced carcinogenesis.

This study is the first to identify EYA2 as a candidate oncogene in cervical carcinogenesis. Next to increased EYA2 expression in hgCIN and carcinomas compared with controls, its functional relevance in HPV-mediated transformation was shown. Silencing of EYA2 in HPV-immortalised cells reduced viability, migration and anchorage-independent growth, substantiating its potential oncogenic role. EYA2 has been suggested to function as a transcriptional co-activator. Increased expression of EYA2 has been seen in prostate, breast, urinary tract and lung cancer.[35,36] Overexpression of EYA2 has been implicated in increased cell proliferation, migration, invasion, transformation and metastasis in mammary carcinoma cells.[26,37] Upregulation of EYA2 has also been shown to promote tumour growth of ovarian cancer cell lines and was shown to be partly related to genomic amplification of its locus in ovarian cancer specimens.[36]

The most frequent focal loss contained hsa-miR-375 and a direct correlation between (focal) loss and reduced expression in hgCIN was shown. Downregulation of hsa-miR-375 expression in a second series of cervical tissues was confirmed, indicating continual decrease with disease progression. To the best of our knowledge we are the first to show a direct correlation between a focal aberration and altered gene expression in the same tissue specimens. Ectopic expression of hsa-miR-375 in SiHa and CaSki cells reduced viability. Cell migration could not be examined due to strong induction of cell death upon ectopic hsa-miR-375 expression. In an independent and parallel study downregulation of this miRNA was reported to be functionally involved in cervical cancer progression[38] and Li *et al*[39] reported decreased expression of hsa-miR-375 in high-grade CIN and SCC. A tumour suppressive function of hsa-miR-375 is further supported by the observed reduced expression in gastric, head and neck, pancreatic, hepatocellular carcinomas and melanomas and the induction of apoptosis and reduced cell viability upon re-expression in gastric, hepatocellular, head and neck, breast cancer as well as melanoma cell lines.[31,40-52] Candidate target genes of hsa-miR-375 include SP1, AEG1 and MTDH.[38,48,49]

Functional studies were restricted to EYA2 and hsa-miR-375 and further mechanistic studies are warranted for the remaining genes validated in the external datasets. Additional validation studies on independent series of high-grade

CIN analysed by high-resolution copy number platforms are required to underline the translational relevance of current findings. Unfortunately, such datasets are not yet available.

The current study was performed on precancerous lesions induced by a viral infection, rather than on advanced tumours with variable and often unknown initiating events previously studied for focal aberrations[14,16-19], thus the genetic chaos is likely modest, enhancing the chance of identifying actual driver genes. In this regard also increased HES1 expression (3q29), restricted to hgCIN, is of particular interest. The mRNA expression pattern is in line with recent findings showing an increase in NOTCH1 expression, an upstream regulator of HES1, in non-tumourigenic HPV-transformed keratinocytes, yet reduced NOTCH1 expression in tumourigenic cervical carcinoma cell lines.[27,53] This is suggestive of a transient induction of NOTCH signaling driving HPV-mediated malignant transformation. Hence, it can be speculated that increased NOTCH signaling, either or not supplemented with the focal gain at 3q29, results in elevated HES1 expression in hgCIN. Genes within the focal aberrations that could not be verified using the external datasets due to their absence await future expression analysis.

Despite the fact that the number of chromosomal aberrations in hgCIN is in part HPV type dependent[13], the present study revealed no correlations between specific focal aberrations and the HPV type present.

To conclude, analysis of focal chromosomal aberrations in hgCIN identified new candidate driver genes in HPV-induced carcinogenesis of the cervix. In addition to demonstrating a direct association between focal loss at 2q35 and decreased expression of hsa-miR-375, functional proof is provided for focal chromosomal aberrations actively contributing to cervical carcinogenesis. EYA2 was identified as a novel candidate oncogene and hsa-miR-375 as a tumour suppressor gene in HPV-mediated transformation.

Reference List

1. zur Hausen H. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. J Natl Cancer Inst 2000; 92: 690-698.

2. Baseman JG and Koutsky LA. The epidemiology of human papillomavirus infections. J Clin Virol 2005; 32 Suppl 1: S16-S24.

3. Snijders PJ, Steenbergen RD, Heideman DA et al. HPV-mediated cervical carcinogenesis: concepts and clinical implications. J Pathol 2006; 208: 152-164.

4. Wentzensen N and von Knebel DM. Biomarkers in cervical cancer screening. Dis Markers 2007; 23: 315-330.

5. Heselmeyer K, Schrock E, du Manoir S et al. Gain of chromosome 3q defines the transition from severe dysplasia to invasive carcinoma of the uterine cervix. Proc Natl Acad Sci U S A 1996; 93: 479-484.

6. Kirchhoff M, Rose H, Petersen BL et al. Comparative genomic hybridization reveals non-random chromosomal aberrations in early preinvasive cervical lesions. Cancer Genet Cytogenet 2001; 129: 47-51.

7. Narayan G, Pulido HA, Koul S et al. Genetic analysis identifies putative tumor suppressor sites at 2q35-q36.1 and 2q36.3-q37.1 involved in cervical cancer progression. Oncogene 2003; 22: 3489-3499.

8. Alazawi W, Pett M, Strauss S et al. Genomic imbalances in 70 snap-frozen cervical squamous intraepithelial lesions: associations with lesion grade, state of the HPV16 E2 gene and clinical outcome. Br J Cancer 2004; 91: 2063-2070.

9. Wilting SM, Steenbergen RD, Tijssen M et al. Chromosomal signatures of a subset of high-grade premalignant cervical lesions closely resemble invasive carcinomas. Cancer Res 2009; 69: 647-655.

10. Bierkens M, Wilting SM, van Wieringen WN et al. Chromosomal profiles of high-grade cervical intraepithelial neoplasia relate to duration of preceding high-risk human papillomavirus infection. Int J Cancer 2012; 131: E579-E585.

11. Ylstra B, van den IJssel P, Carvalho B et al. BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). Nucleic Acids Res 2006; 34: 445-450.

12. Wilting SM, Smeets SJ, Snijders PJ et al. Genomic profiling identifies common HPV-associated chromosomal alterations in squamous cell carcinomas of cervix and head and neck. BMC Med Genomics 2009; 2: 32.

13. Bierkens M, Wilting SM, van Wieringen WN et al. HPV type-related chromosomal profiles in high-grade cervical intraepithelial neoplasia. BMC Cancer 2012; 12: 36.

14. Beroukhim R, Mermel CH, Porter D et al. The landscape of somatic copy-number alteration across human cancers. Nature 2010; 463: 899-905.

15. Krijgsman O, Israeli D, Haan JC et al. CGH arrays compared for DNA isolated from formalin-fixed, paraffin-embedded material. Genes Chromosomes Cancer 2012; 51: 344-352.

16. Weir BA, Woo MS, Getz G et al. Characterizing the cancer genome in lung adenocarcinoma. Nature 2007; 450: 893-898.

17. Leary RJ, Lin JC, Cummins J et al. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. Proc Natl Acad Sci U S A 2008; 105: 16224-16229.

18. Brosens RP, Haan JC, Carvalho B et al. Candidate driver genes in focal chromosomal aberrations of stage II colon cancer. J Pathol 2010; 221: 411-424.

19. Varambally S, Cao Q, Mani RS et al. Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science 2008; 322: 1695-1699.

20. Bulkmans NW, Berkhof J, Rozendaal L et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. Lancet 2007; 370: 1764-1772.

21. Steenbergen RD, Walboomers JM, Meijer CJ et al. Transition of human papillomavirus type 16 and 18 transfected human foreskin keratinocytes towards immortality: activation of telomerase and allele losses at 3p, 10p, 11q and/or 18q. Oncogene 1996; 13: 1249-1257.

22. Steenbergen RD, Kramer D, Braakhuis BJ et al. TSLC1 gene silencing in cervical cancer cell lines and cervical neoplasia. J Natl Cancer Inst 2004; 96: 294-305.

23. Buffart TE, Tijssen M, Krugers T et al. DNA quality assessment for array CGH by isothermal whole genome amplification. Cell Oncol 2007; 29: 351-359.

24. Buffart TE, Israeli D, Tijssen M et al. Across array comparative genomic hybridization: a strategy to reduce reference channel hybridizations. Genes Chromosomes Cancer 2008; 47: 994-1004.

25. van de Wiel MA, Picard F, van Wieringen WN et al. Preprocessing and downstream analysis of microarray DNA copy number profiles. Brief Bioinform 2011; 12: 10-21.

26. Farabaugh SM, Micalizzi DS, Jedlicka P et al. Eya2 is required to mediate the pro-metastatic functions of Six1 via the induction of TGF-beta signaling, epithelial-mesenchymal transition, and cancer stem cell properties. Oncogene 2012; 31: 552-562.

27. Henken FE, De-Castro AJ, Rosl F et al. The functional role of Notch signaling in HPV-mediated transformation is dose-dependent and linked to AP-1 alterations. Cell Oncol (Dordr) 2012; 35: 77-84.

28. Schmittgen TD and Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 2008; 3: 1101-1108.

29. Scotto L, Narayan G, Nandula SV et al. Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression. Genes Chromosomes Cancer 2008; 47: 755-765.

30. Zhai Y, Kuick R, Nan B et al. Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. Cancer Res 2007; 67: 10163-10172.

31. Mazar J, DeBlasio D, Govindarajan SS et al. Epigenetic regulation of microRNA-375 and its role in melanoma development in humans. FEBS Lett 2011; 585: 2467-2476.

32. Futreal PA, Coin L, Marshall M et al. A census of human cancer genes. Nat Rev Cancer 2004; 4: 177-183.

33. Conrad DF, Pinto D, Redon R et al. Origins and functional impact of copy number variation in the human genome. Nature 2010; 464: 704-712.

34. Wilting SM, Snijders PJ, Meijer GA et al. Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. J Pathol 2006; 209: 220-230.

35. Su AI, Welsh JB, Sapinoso LM et al. Molecular classification of human carcinomas by use of gene expression signatures. Cancer Res 2001; 61: 7388-7393.

36. Zhang L, Yang N, Huang J et al. Transcriptional coactivator Drosophila eyes absent homologue 2 is up-regulated in epithelial ovarian cancer and promotes tumor growth. Cancer Res 2005; 65: 925-932.

37. Pandey RN, Rani R, Yeo EJ et al. The Eyes Absent phosphatase-transactivator proteins promote proliferation, transformation, migration, and invasion of tumor cells. Oncogene 2010; 29: 3715-3722.

38. Wang F, Li Y, Zhou J et al. miR-375 is downregulated in squamous cervical cancer and inhibits cell migration and invasion via targeting transcription factor SP1. Am J Pathol 2011; 179: 2580-2588.

39. Li Y, Wang F, Xu J et al. Progressive miRNA expression profiles in cervical carcinogenesis and identification of HPV-related target genes for miR-29. J Pathol 2011; 224: 484-495.

40. Szafranska AE, Davison TS, John J et al. MicroRNA expression alterations are linked to tumorigenesis and non-neoplastic processes in pancreatic ductal adenocarcinoma. Oncogene 2007; 26: 4442-4452.

41. Ladeiro Y, Couchy G, Balabaud C et al. MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. Hepatology 2008; 47: 1955-1963.

42. Avissar M, Christensen BC, Kelsey KT et al. MicroRNA expression ratio is predictive of head and neck squamous cell carcinoma. Clin Cancer Res 2009; 15: 2850-2855.

43. de Souza Rocha SP, Breiling A, Gupta N et al. Epigenetically deregulated microRNA-375 is involved in a positive feedback loop with estrogen receptor alpha in breast cancer cells. Cancer Res 2010; 70: 9175-9184.

44. Ding L, Xu Y, Zhang W et al. MiR-375 frequently downregulated in gastric cancer inhibits cell proliferation by targeting JAK2. Cell Res 2010; 20: 784-793.

45. Hui AB, Lenarduzzi M, Krushel T et al. Comprehensive MicroRNA profiling for head and neck squamous cell carcinomas. Clin Cancer Res 2010; 16: 1129-1139.

46. Liu AM, Poon RT and Luk JM. MicroRNA-375 targets Hippo-signaling effector YAP in liver cancer and inhibits tumor properties. Biochem Biophys Res Commun 2010; 394: 623-627.

47. Tsukamoto Y, Nakada C, Noguchi T et al. MicroRNA-375 is downregulated in gastric carcinomas and regulates cell survival by targeting PDK1 and 14-3-3zeta. Cancer Res 2010; 70: 2339-2349.

48. He XX, Chang Y, Meng FY et al. MicroRNA-375 targets AEG-1 in hepatocellular carcinoma and suppresses liver cancer cell growth in vitro and in vivo. Oncogene 2012; 31: 3357-69.

49. Hui AB, Bruce JP, Alajez NM et al. Significance of Dysregulated Metadherin and MicroRNA-375 in Head and Neck Cancer. Clin Cancer Res 2011; 17: 7539-7550.

50. LaConti JJ, Shivapurkar N, Preet A et al. Tissue and serum microRNAs in the Kras(G12D) transgenic animal model and in patients with pancreatic cancer. PLoS ONE 2011; 6: e20687.

51. Nohata N, Hanazawa T, Kikkawa N et al. Tumor suppressive microRNA-375 regulates oncogene AEG-1/MTDH in head and neck squamous cell carcinoma (HNSCC). J Hum Genet 2011; 56: 595-601.

52. Wiklund ED, Gao S, Hulf T et al. MicroRNA Alterations and Associated Aberrant DNA Methylation Patterns across Multiple Sample Types in Oral Squamous Cell Carcinoma. PLoS ONE 2011; 6: e27840.

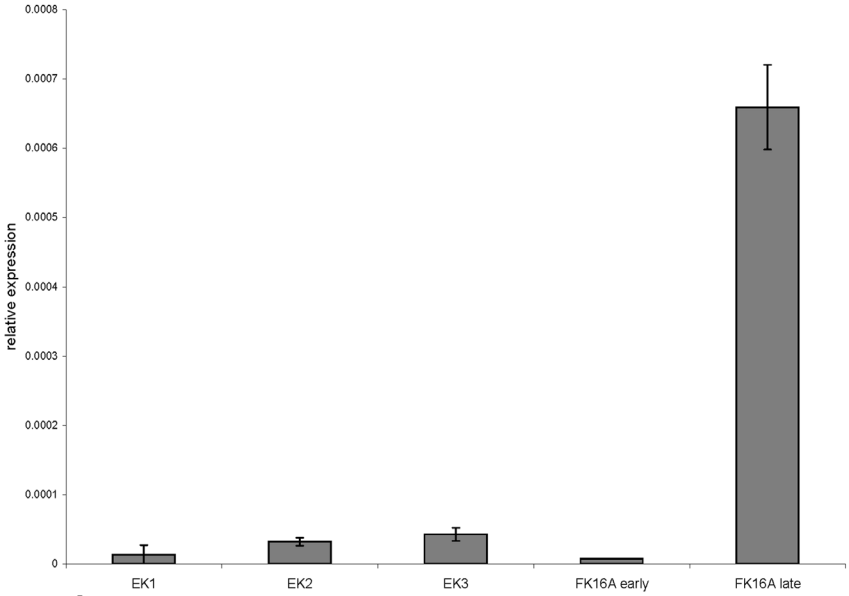
53. de Wilde J, De-Castro AJ, Snijders PJ et al. Alterations in AP-1 and AP-1 regulatory genes during HPV-induced carcinogenesis. Cell Oncol 2008; 30: 77-87.

Supplementary Table I: Genes within the high frequency regions (HFR) of the identified focal chromosomal aberrations.

Chromosome	Start_HFR	End_HFR	Size of HFR	Frequency gain	Frequency loss	Frequency focal gain	Frequency focal loss	Genes
1	76347717	76772053	0.42	20	2	0	2	ST6GALNAC3
1	177742624	178392811	0.65	17	0	2	0	QSOX1, TOR1AIP1, CEP350, TDRD5, FAM163A, TOR1AIP2, NPHS2, C1orf125
1	235193904	235372299	0.18	18	2	0	2	MT1P2, RYR2
2	42264744	42382034	0.12	3	5	2	0	EML4
2	44996756	45136143	0.14	1	7	0	3	SIX2, SIX3
2	71650128	72054559	0.4	1	8	0	4	DYSF
2	176673584	176745707	0.07	0	7	0	3	MIRN10B, RPLP1P4, HOXD3, HOXD4, HOXD8
2	219554248	219604836	0.05	0	10	0	5	CCDC108, CRYBA2, FEV, MIRN375
2	220239129	222152032	1.91	2	7	2	0	EPHA4
3	16930609	17059508	0.13	7	2	3	0	PLCL2
3	24491603	24630003	0.14	16	2	11	0	THRB
3	42682419	42725965	0.04	6	4	0	2	CCDC13, HHATL, KBTBD5, ZBTB47
3	51990183	52016719	0.03	7	2	2	0	RPL29P11, RPL29P12, RPL29P9, RPL29P26, RPL29, ACY1, ABHD14A
3	194023770	195677189	1.65	27	0	2	0	ATP13A3, GP5, LRRC15, CPN2, HES1, OPA1, ATP13A4, ATP13A5, HRASLS, C3orf59
4	41832776	42663376	0.83	0	12	0	3	GRXCR1, ATP8A1, SHISA3, BEND4
4	84066493	84390039	0.32	4	7	2	0	PLAC8, COPS4, LIN54
5	63279090	63676509	0.4	3	9	0	5	RNF180, HTR1A
5	90160596	90712873	0.55	6	4	3	0	ARRDC3, GPR98
5	95214303	96390341	1.18	6	4	3	0	MIRN583, , LNPEP, ERAP2, ERAP1, CAST, PCSK1, ELL2
5	106704954	107003356	0.3	6	4	3	0	EFNA5
5	162804051	162850776	0.05	4	4	2	0	HMMR, NUDCD2, CCNG1
5	178601902	178695617	0.09	2	7	0	3	ADAMTS2
6	7525022	7863330	0.34	2	7	0	2	TXNDC5, MUTED, BMP6, SNRNP48, DSP
6	26212458	26333355	0.12	5	5	3	0	HIST1H1PS1, RPS10P1, HIST1H2BG, HIST1H2AE, HIST1H4E, HIST1H2BF, HIST1H4D, HIST1H3D, HIST1H2AD, HIST1H2BE, HIST1H2BD, HIST1H1E, HIST1H2AC, HIST1H2BC, HIST1H1T, HIST1H4C, HIST1H4A
6	27882725	27947673	0.06	5	5	3	0	HIST1H4PS1, HIST1H2BPS2, HIST1H3I, HIST1H1B, HIST1H2AL, HIST1H2BN, HIST1H2AK, HIST1H4K, HIST1H4J, HIST1H2BM, HIST1H2AJ, HIST1H3H, HIST1H2AI, HIST1H2BL
6	30871594	30952195	0.08	2	5	2	0	
6	40624689	41748197	1.12	0	10	0	5	MDFI, FOXP4, NCR2, TREM1, TREML4, TREML3, TREML2P, TREML2, TREM2, TREML1, NFYA, C6orf130, APOBEC2, BZRPL1, UNC5CL, LRFN2
7	69325536	69441320	0.12	1	4	2	0	AUTS2
7	73105632	73429720	0.32	0	8	0	3	CLIP2, RFC2, LAT2, EIF4H, LIMK1, ELN, MIRN590
7	79590368	80209836	0.62	2	5	2	0	SEMA3C, CD36, GNAT3, GNAI1
7	108206652	109055550	0.85	0	7	0	2	C7orf66
8	33662400	34715128	1.05	5	10	0	4	RPL10AP3
8	82359733	82664878	0.31	13	1	6	0	FABP12, FABP4, FABP9, PMP2
8	98196003	98477825	0.28	8	3	0	2	TSPYL5

9	4730391	5289996	0.56	2	3	0	2	MIRN101-2, RLN2, INSL4, INSL6, IGHEP2, JAK2, RCL1, AK3
9	11808290	11997106	0.19	5	1	3	0	
9	34500054	34681047	0.18	2	3	0	2	CCL19, , CCL27, IL11RA, GALT, SIGMAR1, ARID3C, DCTN3, C9orf23, CNTFR, ENHO, DNAI1
9	130810175	132952292	2.14	3	2	0	2	LAMC3, QRFP, FIBCD1, ABL1, EXOSC2, PRDM12, FUBP3, ASS1, HMCN2, FREQ, GPR107, FNBP1, USP20, C9orf78, TOR1B, TOR1A, PTGES, PRRX2, ASB6, METTL11A, C9orf50, C9orf106, IER5L, PPP2R4, CRAT, DOLPP1, FAM73B, SH3GLB2
10	17931708	18287955	0.36	0	2	3	0	SLC39A12, MRC1, FAM23B, MRC1L1, MIRN511-2, MIRN511-1
10	72671625	73417685	0.75	0	4	0	2	CHST3, PSAP, C10orf54, C10orf105, CDH23, SLC29A3, UNC5B
10	105742742	105873884	0.13	4	2	4	0	MIRN936, C10orf78, COL17A1, SLK
11	26962636	27474773	0.51	5	3	4	0	LIN7C, LGR4, CCDC34, BBOX1, FIBIN
11	34571857	34890152	0.32	4	2	3	0	APIP, EHF
11	45950245	46033768	0.08	6	3	5	0	PHF21A
11	76276789	76374391	0.1	4	2	2	0	ACER3
11	119703907	119865059	0.16	3	6	2	0	ARHGEF12, TMEM136
12	4539937	5025977	0.49	2	10	0	2	KCNA5, KCNA1, KCNA6, GALNT8, NDUFA9, AKAP3, DYRK4
12	50177525	50709085	0.53	1	4	0	2	GRASP, ACVR1B, ACVRL1, ANKRD33, FIGNL2, SCN8A, SLC4A8
12	68289516	68419921	0.13	2	4	0	2	RAB3IP, BEST3, LRRC10
12	101817805	101877046	0.06	2	10	0	8	ASCL1, PAH
12	103762476	104427129	0.66	4	2	2	0	APPL2, KIAA1033, ALDH1L2, C12orf45, SLC41A2
13	27194657	28535882	1.34	1	4	0	2	KIAA0774, SLC46A3, POMP, FLT1, PAN3, FLT3, PRHOXNB, CDX2, ATP5EP2, PDX1, GSX1, EEF1AL1
13	52481172	52735267	0.25	3	2	2	0	OLFM4
13	72512160	73181475	0.67	10	2	10	0	KLF12, KLF5
13	85286335	86020507	0.73	6	2	5	0	
13	94533409	94611536	0.08	4	2	2	0	ABCC4
13	103025470	105648392	2.62	3	2	2	0	DAOA
16	52574827	53503466	0.93	12	1	8	0	IRX3, FTO
16	71473267	71718806	0.25	5	3	0	2	C16orf47, ZFHX3
16	73341676	73631283	0.29	9	1	4	0	ZNRF1, WDR59, FA2H
17	28368756	29488485	1.12	0	7	0	2	ACCN1
17	43994277	44039891	0.05	2	5	2	0	MIRN10A, HOXB7, HOXB6, HOXB5, HOXB4, HOXB3
17	50396404	50545965	0.15	2	5	2	0	STXBP4, COX11
18	13185387	13289033	0.1	1	6	0	2	C18orf1
18	26991419	27298126	0.31	11	3	9	0	DSG3, DSG4, DSG1, DSC1
18	33122866	33412930	0.29	2	5	0	2	BRUNOL4
18	44483911	44805874	0.32	2	5	0	2	SMAD7, KIAA0427
18	75235681	75427825	0.19	1	5	0	2	NFATC1, ATP9B
20	843253	974895	0.13	12	3	0	2	RSPO4, ANGPT4
20	31552311	31700561	0.15	14	0	2	0	CBFA2T2
20	44945757	45671015	0.73	14	0	4	0	NCOA3, ZMYND8, EYA2, RPL35AP
21	16508005	18091291	1.58	5	1	3	0	MIRN99A, C21orf91, BTG3, CXADR, C21orf34, MIRN125B2, MIRNLET7C
22	35495013	35543776	0.05	3	6	0	3	PVALB, RABL4
22	44467513	44614200	0.15	6	4	4	0	ATXN10

Supplementary Figure I



Supplementary Figure I: EYA2 mRNA expression in primary foreskin keratinocytes (EK) and HPV-transformed keratinocytes (FK16A) early/late passage. The late passage cell line was selected for siRNA experiments.